



## An improved TROPOMI tropospheric HCHO retrieval over China

Wenjing Su<sup>1\*</sup>, Cheng Liu<sup>2,3,4,5,6,\*</sup>, Ka Lok Chan<sup>7</sup>, Qihou Hu<sup>2</sup>, Haoran Liu<sup>4</sup>, Xiangguang Ji<sup>2,8</sup>, Yizhi Zhu<sup>2</sup>, Ting Liu<sup>1</sup>, Chengxin Zhang<sup>1</sup>, Yujia Chen<sup>2</sup>, Jianguo Liu<sup>2</sup>

<sup>1</sup>School of Earth and Space Sciences, University of Science and Technology of China, Hefei 230026, China

5 <sup>2</sup>Key Lab of Environmental Optics and Technology, Anhui Institute of Optics and Fine Mechanics, Hefei Institutes of Physical Science, Chinese Academy of Sciences, Hefei 230031, China

<sup>3</sup>Center for Excellence in Regional Atmospheric Environment, Institute of Urban Environment, Chinese Academy of Sciences, Xiamen 361021, China

10 <sup>4</sup>Department of Precision Machinery and Precision Instrumentation, University of Science and Technology of China, Hefei 230027, China

<sup>5</sup>Key Laboratory of Precision Scientific Instrumentation of Anhui Higher Education Institutes, University of Science and Technology of China, Hefei 230027, China

<sup>6</sup>Anhui Province Key Laboratory of Polar Environment and Global Change, USTC, Hefei 230026, China

<sup>7</sup>Remote Sensing Technology Institute (IMF), German Aerospace Center (DLR), Oberpfaffenhofen, Germany

15 <sup>8</sup>School of Environmental Science and Optoelectronic Technology, University of Science and Technology of China, Hefei 230026, China

*Corresponding author:* Cheng Liu (chliu81@ustc.edu.cn) and Ka Lok Chan (ka.chan@dlr.de)

**Abstract.** We present the improved retrieval of Tropospheric Monitoring Instrument (TROPOMI) tropospheric formaldehyde (HCHO) over China. The new retrieval optimizes the slant column density (SCD) retrieval and air mass factor (AMF) calculation for TROPOMI observations of HCHO over China. HCHO SCDs are retrieved using the basic optical differential spectroscopy (BOAS) technique, while AMFs are calculated with a priori HCHO profile from a higher resolution regional chemistry transport model. Compared to the operational product, the new TROPOMI HCHO retrieval shows better agreement with the ground based Multi-AXis Differential Optical Absorption Spectroscopy (MAX-DOAS) measurements in China. The operational product in general overestimates HCHO VCDs by 14.01 %, while the improved HCHO only shows an  
20 underestimation of 3.67 %. The improvements are mainly related to the AMF calculation with higher resolution a priori profile (61.11 %), while the SCD retrieval only shows a minor effect of 0.15 %. The improved HCHO is also used to investigate the



spatial-temporal characteristic of HCHO over China. The result shows that HCHO VCDs reach maximum in summer and minimum in winter. High HCHO VCDs mainly located over populated areas, i.e., Sichuan Basin, Central and Eastern China, indicating a significant contribution of anthropogenic emissions. The result indicates the improved TROPOMI HCHO is more  
30 suitable for the analysis of regional and city scale pollution in China.

## 1 Introduction

Formaldehyde (HCHO) is an important trace gas playing a crucial role in atmospheric chemistry process. Hydroxyl radicals ( $\text{HO}_x = \text{OH} + \text{HO}_2$ ) produced from HCHO photolysis show a strong influence on the oxidative capacity of the atmosphere (Li et al., 2011; Xue et al., 2016), and contribute to the formation of secondary organic aerosol (SOA) (Jang and Kamens, 2001).  
35 Atmospheric HCHO can be emitted from both primary and secondary sources. Primary emissions include industrial sources, vehicle emissions (Wei et al., 2008), biogenic emissions from vegetation and biomass burning (Bauwens et al., 2016). The contribution of secondary sources to ambient HCHO sources is in general much higher than the primary emissions, especially in summer (Su et al., 2019). Secondary HCHO is formed through the oxidation of almost all volatile organic compounds (VOCs). Therefore, it is usually regarded as an indicator of VOCs and being used to analyze sensitivity regimes of ozone ( $\text{O}_3$ )  
40 formation (Martin et al., 2004; Choi et al., 2012; Jin and Holloway, 2015; Liu et al., 2016). Accurate HCHO measurement is important for the investigation of HCHO and  $\text{O}_3$  interactions and for the understanding of atmospheric chemistry process. Satellite observations provide indispensable information on HCHO spatial distribution. Satellite observations of HCHO have been conducted since 1996 with a series of satellite borne instruments (Bovensmann et al., 1999; Martin et al., 2004; De Smedt et al., 2012; Barkley et al., 2013; González Abad et al., 2015; González Abad et al., 2016). TROPospheric Monitoring  
45 Instrument (TROPOMI) was launched on 13 October 2017. Compared its predecessor satellite instruments, TROPOMI provides HCHO observations with a much higher spatial resolution with daily global coverage. The high resolution satellite measurement enables us to analyze the finer scale spatio-temporal characteristic of HCHO. In this study, we have improved the TROPOMI HCHO over China by optimizing the slant column density (SCD) retrieval and the air mass factor (AMF) calculation. Different from the operational product, our retrieval uses the basic optical differential  
50 spectroscopy (BOAS) technique for the SCD retrieval. This technique has been reported featured with lower fitting



uncertainties compared to the standard differential optical absorption spectroscopy (DOAS) method (Chance and Kurosu, 2003). In addition, the AMF calculation is improved by using a priori profile from higher resolution regional model. It takes the fine scale pollution into account and the result is expected to be more realistic for the investigation of spatio-temporal variation of HCHO over China.

55 The paper is organized as follows. Section 2 describes all data sets used in this study. Section 3 presents the improved TROPOMI HCHO retrieval algorithm over China. The comparisons to operational TROPOMI HCHO product and ground based Multi-AXis Differential Optical Absorption Spectroscopy (MAX-DOAS) measurements in China are shown in Section 4. Finally, the summary and conclusion are drawn in Section 5.

## 2 Data Sets

### 60 2.1 The TROPOMI instrument

TROPOMI is onboard the Sentinel-5 Precursor (S5P) satellite. The S5P satellite orbits on the near-polar sun-synchronous orbit at an attitude of 824 km with a 17 days repeat cycle and equator-crossing time of 13:30 h local solar time (LST) on the ascending node. A scanning swath of TROPOMI covers a width of 2600 km providing daily global coverage. TROPOMI has four two-dimensional spectrometers covering wavelengths from 270 to 2385 nm, which is divided in eight wavelength bands.

65 The third wavelength band from 320 to 405 nm is used for HCHO retrieval. The spectral resolution of this wavelength band is about 0.5 nm (FWHM). The spatial resolution of the instrument is  $3.5 \times 7 \text{ km}^2$  in nadir.

### 2.2 Operational TROPOMI HCHO product

The operational TROPOMI HCHO product is jointly developed at the German Aerospace Center (DLR) and Royal Belgian Institute for Space Aeronomy (BIRA) (available at <https://scihub.copernicus.eu/>, last access: 22 May 2019). The operational  
70 product is used to compare to the improved TROPOMI HCHO data set over China. The operational product retrieves HCHO columns using the DOAS spectral fitting technique in the wavelength range of 328.5-359 nm with earthshine radiance over the remote Pacific as reference. The residual HCHO signal in the reference spectra is corrected by reference sector correction. The conversion of SCD to VCD uses the AMF approach. A priori HCHO profile is taken from the global chemistry transport

model Tracer Model 5 (TM5-MP) with a spatial resolution of  $1^\circ \times 1^\circ$ . Detail of S5P operational HCHO algorithm can be found  
75 in De Smedt et al. (2018).

### 2.3 WRF-Chem model

In our retrieval, the regional Weather Research and Forecasting model (WRF-Chem) is used to simulate a priori HCHO profile. Compared to TM5-MP, the regional WRF-Chem simulation has a higher spatial resolution of  $20 \times 20 \text{ km}^2$  and more up to date emission inventory of China. WRF-Chem simulates air pollutions from July 2018 to July 2019 with 44 vertical layers  
80 from the ground level up to 50 hPa. The initial and boundary conditions of the meteorological field for simulation are taken from the National Centers for Environmental Prediction (NCEP) 6-hour Final Operational Global (FNL) reanalysis data with spatial and temporal resolutions of  $1^\circ \times 1^\circ$  and 6-hour. The CBMZ (Carbon-Bond Mechanism version Z) photochemical mechanism combined with the MOSAIC (Model for Simulating Aerosol Interactions and Chemistry) aerosol model was used to simulate the chemical process in the atmosphere. The anthropogenic and biogenic emissions are obtained from The Multi-  
85 resolution Emission Inventory for China (MEIC) and the Model of Emissions of Gases and Aerosols from Nature (MEGAN) (Guenther et al., 2006), respectively.

### 2.4 MAX-DOAS HCHO measurements

MAX-DOAS HCHO measurements are used to validate TROPOMI HCHO observations in this study. The MAX-DOAS measurements are performed at three sites in Beijing (Fig. 1), including one urban and two suburban sites. One suburban site  
90 is located in Nancheng (NC) in the southern side of Beijing, while another suburban site is located at the University of Chinese Academy of Sciences (UCAS) in the northeastern side of Beijing. The urban site is located in Chinese academy of meteorological sciences (CAMS). The MAX-DOAS instrument consists of a scanning telescope, a stepping motor controlling the viewing direction of the telescope and a spectrometer with Hamamatsu backthinned charge-coupled device (CCD) detector. The spectrometer measures scattered sunlight in the spectral range of 300 - 505 nm. Details of the MAX-DOAS measurement  
95 setup are shown in Table 1. Scattered sunlight spectra measured by MAX-DOAS are recoded and analyzed using DOAS Intelligent System (DOASIS) spectral fitting software (Kraus, 2006). Details of the MAX-DOAS HCHO DCSD retrieval



settings are listed in Table S1. HCHO vertical profiles are retrieved using Munich Multiple wavelength MAX-DOAS retrieval algorithm ( $M^3$ ) (Chan et al., 2018). The algorithm is developed based on the optimal estimation method (Rodgers, 2000) and utilizes the radiative transfer model LibRadTran (Emde et al., 2016) as the forward model. In the MAX-DOAS profile retrieval, the lowest 1 km is divided into 10 layers with the thickness of each layer of 100 m, while the thickness of the layers between 1 km and 3 km is set to 200 m. Details of MAX-DOAS HCHO profile retrieval can be found in (Chan et al., 2018; Chan et al., 2019).

### 3 Improved HCHO retrieval algorithm

The retrieval of HCHO vertical column densities (VCDs) from TROPOMI observations can be separated into three major steps. The first step is the retrieval of HCHO SCD. The second step is the reference sector correction. The final step is the conversion of corrected SCD to VCD using AMF. Details of the improved HCHO retrieval algorithm are presented in the following.

#### 3.1 HCHO SCD retrieval

##### 3.1.1 Wavelength calibration

Measured solar irradiances and earthshine radiances are often slightly misaligned with the on ground calibration due to temperature variation of the spectrograph. To achieve accurate radiance fitting, it is necessary to calibrate the wavelength mapping before further processing. Irradiance and radiance wavelength are calibrated using a high resolution solar spectrum with accuracy of 0.001 nm (Chance and Kurucz, 2010). The high resolution solar spectrum is convolved with the preflight instrument slit function to calculate the solar spectrum with the instrument resolution. This process can be described by Eq. 1:

$$I_r = I_0^h \otimes S(\lambda + \Delta\lambda) \times P_s^m(\lambda) + P_b^m(\lambda) \quad (1)$$

Where  $I_r$  is the calculated solar spectrum with the instrument resolution.  $I_0^h$  is the high resolution solar spectrum.  $\lambda$  represents the wavelength.  $\Delta\lambda$  indicates the wavelength shift parameter. The symbol  $\otimes$  represents convolution procedure.  $P_s^m(\lambda)$  and  $P_b^m(\lambda)$  are scaling and baseline polynomial, respectively, which account for the low-frequency structures of the



measured spectra. The polynomial is calculated as:

$$120 \quad P^m(\lambda) = \sum_{i=0}^m C_i(\lambda - \lambda_{avg})^i \quad (2)$$

Where  $\lambda_{avg}$  is the center wavelength of the fitting window.  $C_i$  is the coefficient of the fitted polynomial and  $m$  is the order of polynomial. Wavelength shift is determined by fitting the  $I_r$  to the measured irradiance using the Gauss-Newton Nonlinear Least Squares (NLLS) method.

### 3.1.2 Radiance fitting

125 HCHO slant column densities (SCDs) are retrieved using the basic optical differential spectroscopy (BOAS) method (Chance, 1998) which has been applied to OMI and OMPS HCHO retrieval (González Abad et al., 2015; González Abad et al., 2016). The BOAS method is based on the direct radiance fit. HCHO SCDs are determined by fitting the simulated and measured radiance using the NLLS method. The simulated radiance  $I_s(\lambda)$  is calculated following the Eq. 3:

$$I_s(\lambda) = [(aI_0(\lambda + \Delta\lambda) + \alpha_r X_r(\lambda)) \times e^{-\sum_j \alpha_j X_j(\lambda)}] \times P_s^m(\lambda) + P_b^m(\lambda) \quad (3)$$

130 Where  $\lambda$  represents the wavelength.  $a$  is the scaling factor of  $I_0$ .  $I_0(\lambda + \Delta\lambda)$  denotes the daily average earthshine radiance over the remote Pacific.  $\Delta\lambda$  indicates the wavelength shift parameter.  $X_r(\lambda)$  refers to the Raman spectrum and  $\alpha_r$  is the fitted coefficient to  $X_r(\lambda)$ .  $X_j(\lambda)$  refer to the fitted SCD and cross section of trace gas  $j$ , respectively. Detail of TROPOMI HCHO SCD retrieval setting is listed in Table 1. Fig. 2 shows the fitted optical depth and residual for two spectra measured on 6 August 2019 over China (orbit 4211). For the spectrum with higher HCHO SCD, the HCHO absorption can be clearly distinguished in the fitting (see Fig. 2a). The HCHO absorption structures are less significant in the other spectrum as the HCHO SCD is reaching the detection limit (see Fig. 2 b). The detection limit of HCHO SCD can be expressed by the ratio of root mean square (RMS) to the peak to peak optical density of HCHO in the fitting window (Schönhardt A, 2008). The RMS of fit residual for TROPOMI observation typically varies from  $5 \times 10^{-4}$  to  $1 \times 10^{-3}$ , corresponding to detection limits of HCHO SCD from  $\sim 7 \times 10^{15}$  to  $\sim 1 \times 10^{16}$  molec cm<sup>-2</sup>.



### 140 3.2 AMF calculation

HCHO SCDs represent the integration of HCHO concentration along the light path. Therefore, it is strongly dependent on the viewing and solar geometries, surface albedo, and the state of the atmosphere (cloud, vertical HCHO profile, pressure, temperature, etc). The AMF approach is used to convert SCD to VCD. AMFs can be calculated using a comprehensive radiative transfer model. For optically thin species, the height dependent sensitivity (box AMF) is insensitive to the vertical profile of  
145 the species. Therefore, tropospheric HCHO AMF is calculated following the approach of Palmer et al. (2001), which can be described as:

$$\text{AMF} = \frac{\int_{z_s}^{z_t} w \times n_a dz}{\int_{z_s}^{z_t} n_a dz} \quad (4)$$

where  $n_a$  is the partial column of the corresponding layer of the a priori profile, which the a priori profile is taken from the regional WRF-Chem simulation.  $z_s$  and  $z_t$  are the altitude of surface and the tropopause, respectively. Box AMF ( $w$ )  
150 represents the height dependent sensitivity of the TROPOMI measurement. The box AMF depends wavelength, solar zenith angle (SZA), viewing zenith angle (VZA), relative azimuth angle (RAA), surface albedo, surface pressure, cloud albedo, cloud fraction and cloud pressure. As the wavelength dependency of box AMF in the HCHO fitting windows rather small (less than 5 % for  $\text{SZA} < 70^\circ$ ), the box AMF can be calculated at a representative wavelength of 340nm. In order to improve the computational efficiency, the box AMF are pre-calculated using the radiative transfer model VLIDORT (version 2.6) with a  
155 number of surface albedos, surface pressures, solar and viewing geometries and stored in a look up table (LUT). An aerosol free U.S. standard atmosphere is assumed in the radiative transfer calculation of box AMF. Box AMF within the LUT is interpolated into each particular observation condition. Linear interpolation is performed in solar and viewing geometries and surface albedo dimensions and a nearest neighbor interpolation is performed in surface pressure dimension. Box AMF at each particular observation condition is then linearly interpolated to the pressure level of the a priori HCHO profile.

160 For partly cloudy pixels, cloud correction is applied using the independent pixel approximation (Martin et al., 2002) in which scattering weights of inhomogeneous scenes are considered as a linear combination of box AMF for cloud-free scenes ( $w_{clear}$ ) and cloudy scene ( $w_{cloud}$ ) following Eq. (5):



$$w(z) = (1 - CF_{iw}) \times w_{clear} + CF_{iw} \times w_{cloud} \quad (5)$$

where  $CF_{iw}$  is the intensity-weighted cloud fraction, defined as:

$$CF_{iw} = \frac{C_f \times I_{cloud}}{(1 - C_f) \times I_{clear} + C_f \times I_{cloud}} \quad (6)$$

Where  $C_f$  is cloud fraction.  $I_{cloud}$  and  $I_{clear}$  are the radiance intensities for cloudy scene and cloud-free scene, respectively. Radiance intensities are pre-calculated using VLIDORT and saved in the LUT with the same setting as box AMF LUT. The surface albedo, surface pressure, cloud fraction, cloud albedo and cloud pressure are obtained from the S5P operational cloud product (Loyola et al., 2018).

### 170 3.3 Reference sector correction of SCDs

Using earthshine radiance over remote Pacific Ocean as reference significantly reduced the across-track striping issue and fitting residual. The striping issue is known to cause by small optical mismatches between radiance and irradiance. Using radiance as reference reduces the fit residual as it already account for the O<sub>3</sub> absorption and Ring effect. Since there are residual HCHO signals in the earthshine reference spectra, therefore, an offset correction has to be applied to the retrieved SCDs.

175 The SCD and AMF for each pixel of the remote Pacific orbit are retrieved using average earthshine radiance over remote Pacific Ocean as reference. The HCHO VCD over remote Pacific Ocean is simulated by GEOS-Chem assuming HCHO over this region is mainly from the oxidation of CH<sub>4</sub>. The simulated HCHO SCD is calculated by multiplying the VCD ( $VCD_s$ ) taken from GEOS-Chem with the corresponding AMF ( $AMF_{Pacific}$ ). The difference between simulated SCD and retrieved SCD ( $SCD_{Pacific}$ ) is the SCD bias caused by residual HCHO signal in reference spectrum. Assuming the SCD correction is latitude  
180 dependent, it is calculated at 500 gridded latitude points (lat) ranging from 90° N to 90° S. The SCD correction for each gridded latitude point is calculated using the median value of SCD biases for all pixels with center latitudes within the given contiguous latitude points of the pacific orbit. The SCD correction can be expressed as the following:

$$Corr(lat) = Median(VCD_s(lat) \times AMF_{Pacific} - SCD_{Pacific}) \quad (7)$$

Assuming the SCD correction is constant in the longitude direction, the SCD correction at 500 gridded latitude points are



185 linearly interpolated to the latitude of each pixel over China. The interpolated SCD correction is then applied on the retrieved  
SCDs ( $SCD_r$ ) to calculate the corrected SCD ( $SCD_{corrected}$ ). This process can be described as follows:

$$SCD_{corrected}(k, m) = SCD_r + Interpolate ( Corr(lat) ) \quad (8)$$

## 4 Results and discussions

### 4.1 Comparison of operational and improved HCHO product

190 To analyze the improvement of the new TROPOMI HCHO product, we compared the new HCHO data set to the operational  
product. The contribution of SCD retrieval and AMF calculation are investigated separately. Both the improved and operation  
data sets are compared to the MAX-DOAS HCHO measurements to quantify the improvements. The normalized mean bias  
(NMB) between satellite and MAX-DOAS data is used as benchmark in the comparison. NMBs between two data sets can be  
calculated following:

$$195 \quad NMB = \frac{\sum_{i=1}^{i=n} (V_T(i) - V_M(i))}{\sum_{i=1}^{i=n} V_M(i)} \times 100\% \quad (9)$$

Where  $V_T(i)$  and  $V_M(i)$  is the average tropospheric HCHO VCD measured by TROPOMI and MAX-DOAS on day  $i$ ,  
respectively.

#### 4.1.1 SCD retrieval

The improved TROPOMI HCHO product is retrieved using the BOAS approach, while the operational product uses the DOAS  
200 technique to retrieve the HCHO SCDs. Fig. 3 shows the TROPOMI HCHO SCDs retrieved by BOAS and DOAS methods.  
Both data sets are filtered for root mean square of spectral fit residual (RMS) smaller than  $10^{-3}$ . HCHO SCDs from both data  
sets show very similar spatial pattern over China with higher values over industrialized region such as, Beijing-Tianjin-Hebei  
region (BTH), Yangtze River Delta (YRD), Pearl River Delta (PRD) and Sichuan Basin (SCB). However, the DOAS SCDs  
show more scattered outliers while the BOAS data set show a smoother appearance. As the anthropogenic VOC emission in  
205 Tibet is small (Wu et al., 2016; Li et al., 2017) and the oxidation of VOCs is a significant source of HCHO, HCHO in this  
region is expected to have small spatial variation. The averaged HCHO SCD of the BOAS retrieval is  $0.42 \times 10^{16}$  molec  $\text{cm}^{-2}$



with a standard deviation (SD) of  $0.38 \times 10^{16}$  molec  $\text{cm}^{-2}$  over Tibet (area indicated with blue line), while the DOAS dataset shows higher averaged column of  $0.63 \times 10^{16}$  molec  $\text{cm}^{-2}$  as well as standard deviation of  $0.43 \times 10^{16}$  molec  $\text{cm}^{-2}$ . Larger standard deviation implies the data set contains more outliers. The DOAS HCHO SCDs are rather noisy over Shaanxi province and its surrounding regions, while BOAS SCDs appear to be much smoother with a significant hotspot over Xi'an (provincial capital of Shaanxi). As the spatial distribution of HCHO is expected to be smooth, less noisy SCDs indicate that the BOAS technique is less sensitive to measurement noise.

The scatter plot of the BOAS and DOAS retrieval of HCHO SCDs over China on 06 August 2018 is in Fig. 4 (a). The result shows BOAS and DOAS retrievals agree well with each other with Pearson correlation coefficient (R) of 0.83. The slope and offset of the linear least squares regression line are 0.93 and  $0.24 \times 10^{16}$  molec  $\text{cm}^{-2}$ , respectively. Averaged SCDs on 06 August 2018 over China retrieved by DOAS and BOAS algorithms are  $0.56 (\pm 0.64) \times 10^{16}$  molec  $\text{cm}^{-2}$  and  $0.34 (\pm 0.57) \times 10^{16}$  molec  $\text{cm}^{-2}$ , respectively. SCDs retrieved by DOAS method are 10.18 % higher ( $0.22 \pm 0.36 \times 10^{16}$  molec  $\text{cm}^{-2}$ ) than the BOAS retrieval. On the other hand, RMS of both methods is very similar. The bias between results using BOAS and DOAS method is mainly related to the difference of reference spectrum, retrieval method and retrieval settings. Corrected DOAS HCHO SCDs in operational product and corrected BOAS HCHO SCDs in our retrieval also show good agreement with Pearson correlation coefficient (R) of 0.81 (Fig. 4 (b)). The slope and intercept of the linear least squares regression line are 0.92 and  $0.05 \times 10^{16}$  molec  $\text{cm}^{-2}$ . Corrected DOAS HCHO SCDs are 32.32 % lower ( $0.02 \pm 0.38 \times 10^{16}$  molec  $\text{cm}^{-2}$ ) than corrected BOAS HCHO SCDs.

To evaluate improvement of SCD retrieval, the control variable method is applied on operational product. We updated operational product by replacing SCDs with the SCDs retrieved with the BOAS technique and the rest parts of the retrieval stay the same. The updated and operational product are compared to the MAX-DOAS observations. The NMBs of the two data sets are shown in Table 3. Changing SCD retrieval method only shows a tiny effect on TROPOMI HCHO VCD retrieval (0.15 %). However, using the BOAS HCHO SCDs reduced the overestimation. Besides, the mean random errors relative to BOAS are about 22 % lower than DOAS at three MAX-DOAS sites. The result indicates the BOAS technique provides a slightly more stable retrieval of HCHO SCDs.



#### 4.1.2 AMF calculation

The improved TROPOMI HCHO over China uses a priori profile information from higher resolution regional WRF-Chem simulation, while the operational product use a priori profile from the global TM5-MP model with  $1^\circ \times 1^\circ$  resolution. Fig.5 shows comparisons of vertical HCHO profiles from MAX-DOAS, WRF-Chem and TM5-MP simulations at CAMS site  
235 in spring (March, April, May), summer (June, July, August), autumn (September, October, November) and winter (December, January, February). A priori HCHO profiles simulated by WRF-Chem are similar to MAX-DOAS measurements, while TM5-MP profiles show a larger difference from the MAX-DOAS measurements. The AMFs calculated with WRF-Chem and TM5-MP HCHO profiles as a priori are shown in Fig. 6. The results show a similar spatial pattern. However, WRF-Chem AMF is mostly higher than the one calculated with TM5-MP HCHO profiles. To evaluate influence of a priori HCHO profile on VCD  
240 retrieval, we updated operational product by keeping the SCDs unchanged and use a priori profile from WRF-Chem for the AMF calculation. The NMBs between the updated and operational HCHO and MAX-DOAS HCHO are shown in Table 3. Using a priori HCHO profiles from the regional WRF-Chem simulation in general reduces HCHO VCDs by 15.49 %. The effect of a priori HCHO profile is much larger than that from SCD retrieval. In addition, using a priori HCHO profile from WRF-Chem reduces the bias between TROPOMI and MAX-DOAS by 61.11 %.

#### 245 4.2 Comparison between HCHO VCDs observed by MAX-DOAS and TROPOMI

To validate TROPOMI tropospheric HCHO VCD, TROPOMI HCHO VCDs are compared to the MAX-DOAS measurements. MAX-DOAS measurements are temporally averaged within  $\pm 1$  h around the TROPOMI overpass time, while TROPOMI pixels within 20 km of the MAX-DOAS site are spatially averaged for comparison. TROPOMI data are filtered for quality assurance value (QA value) larger than 0.5 and intensity-weighted cloud fraction smaller than 0.3.

250 The left panels in Fig.7 show the comparison of TROPOMI HCHO VCDs from our retrieval and operational product to the MAX-DOAS HCHO VCDs at three MAX-DOAS sites in Beijing. Our retrieval shows better agreement with MAX-DOAS tropospheric HCHO VCD compared to the operational product. The Pearson correlation coefficients (R) of our retrieval are 0.74, 0.8 and 0.78 at CAMS, UCAS and NC sites, respectively. The corresponding R values of operational product are 0.74, 0.71 and 0.76, which are slightly lower than our retrieval. Besides, the slopes of linear least squares regression lines of our



255 retrieval are 0.57, 0.52 and 0.58 at CAMS, UCAS and NC sites. The corresponding values for the operational retrieval are 0.47, 0.46, and 0.51. In addition, the offsets of linear least squares regression lines of our retrieval are 0.58, 0.53 and  $0.48 \times 10^{16}$  molec  $\text{cm}^{-2}$  at three MAX-DOAS sites, with the corresponding values for operational product of 0.84, 0.85 and  $0.76 \times 10^{16}$  molec  $\text{cm}^{-2}$ . The agreement between our retrieval and MAX-DOAS observations is in general better than the operational product. Moreover, the absolute NMBs of our retrieval relative to MAX-DOAS HCHO are about 66.20 % smaller than that of

260 operational product (Table 3). The results also indicate the operational product overestimated HCHO VCD by 14.01 %, while our retrieval shows a small underestimation of 3.67 %. These results suggest that our retrieval is better than the operational product both in urban and suburban in China. In order to investigate the influence of HCHO a priori profile in the satellite retrieval, we used the MAX-DOAS HCHO profiles as a priori in the TROPOMI AMF calculation. The results show an improvement of correlation and absolute agreements. The Pearson correlation coefficient (R) increases by about 0.07. The

265 linear fitting slopes increase by 0.23 and the offsets decrease by  $0.18 \times 10^{16}$  molec  $\text{cm}^{-2}$ .

The right panels of Fig. 7 show the time series of HCHO VCDs from MAX-DOAS, our retrieval and operational product. HCHO VCDs from our retrieval, operational product and MAX-DOAS show a similar temporal variation pattern, reaching maximum value in summer and minimum value in winter, while the seasonal variation is less significant from the operational product. Both our retrieval and operational product underestimate HCHO VCD in summer. As shown in Fig. 5(b), WRF-Chem

270 and TM5-MP model both underestimate HCHO concentration in the lower troposphere in summer and result in an overestimation of AMFs. The vertical HCHO profiles simulated by WRF-Chem model are similar to the one measured by MAX-DOAS in summer, while the TM5-MP profiles show larger difference to the MAX-DOAS measurements. The underestimation of our retrieval (9.96 %) is slightly less than the operational product (10.88 %) during summer. With lower HCHO columns ( $< 1 \times 10^{16}$  molec  $\text{cm}^{-2}$ ) in winter, our retrieval and operational product both overestimate HCHO VCD and the

275 overestimation of our retrieval (43.24 %) is smaller than operational product (89.57 %). Compared to other seasons, large differences between TOROPOMI and MAX-DOAS HCHO measurements are observed in winter. Using MAX-DOAS HCHO profile as a priori in the satellite retrieval reduces the underestimation in summer and the overestimation in winter in our retrieval. The remaining overestimation is mainly related to a portion of HCHO above 3 km where MAX-DOAS is not sensitive.



### 4.3 Spatial-temporal characteristic of HCHO over China

280 Fig. 8 shows the spatial distribution of HCHO VCDs over China in spring, summer, autumn and winter. HCHO VCDs reach highest values in summer and lowest values in winter over China. This is due to higher biogenic emissions (Guenther et al., 1995; Li and Xie, 2014; Jiang et al., 2019) and secondary formation in summer (Wang et al., 2016; Su et al., 2019). The average HCHO VCDs over China in autumn 2018, winter 2018, spring 2019 and summer 2019 are 0.93, 0.91, 0.98 and  $1.21 \times 10^{16}$  molec  $\text{cm}^{-2}$ , respectively (Table S2). High HCHO VCDs are mainly occurred in central and eastern China and SCB. Lower  
285 HCHO values are observed in the western China. Hotspots can be observed in summer over BTH, YRD, PRD, Shandong province, Henan province, Wuhan (Hubei's provincial capital), SCB and cities along Fen nutrient-laden valley in Shaanxi and Shanxi provinces. These hotspot patterns are strongly correlated to the population density and industrial emission pattern indicating a significant anthropogenic contribution. The annual average HCHO VCDs in BTH, YRD and PRD are similar. Among the provincial administrative region of China, the annual average HCHO VCD is highest in Tianjin and smallest in  
290 Qinghai and Tibet (Table S2).

### 5 Conclusions

In this paper, we present an improved TROPOMI HCHO retrieval over China. We mainly improve the retrieval in two aspects, SCD and AMF calculations. The improved HCHO uses the BOAS technique for the retrieval of HCHO SCDs. Compared to the DOAS retrieval technique used in the operational product, results retrieved with the BOAS technique are less noisy while  
295 the spatial pattern of HCHO remains similar. SCDs retrieval with both BOAS and DOAS correlate well with each other with a correlation coefficient of 0.83 and a linear fitting slope of 0.93.

We also compared the improved TROPOMI HCHO VCDs to MAX-DOAS observations at three MAX-DOAS sites in Beijing. The improved HCHO shows better correlation with MAX-DOAS HCHO VCDs compared to the operational product. In addition, the bias between the improved TROPOMI HCHO VCDs and MAX-DOAS observations is 66.20 % smaller than  
300 operational product. These results suggest our retrieval can better capture the fine scale variation of HCHO in both urban and suburban regions. The improved TROPOMI HCHO VCDs and MAX-DOAS show a similar temporal variation pattern with



higher values in summer and lower values in winter, while the operational product shows less pronounced seasonal pattern. The improved and operational products both underestimate HCHO VCDs in summer while overestimate HCHO VCDs in winter.

305 The analysis of HCHO spatial distribution shows higher values over SCB, Central and Eastern China and SCB. Lower HCHO VCDs can be observed over western China. In summer, hotspots can be observed over BTH, YRD, PRD, Shandong province, Henan province, Wuhan (Hubei's provincial capital), SCB and cities along Fen nutrient-laden valley in Shaanxi and Shanxi provinces. These spatial patterns indicate significant anthropogenic contributions of atmospheric HCHO and the improved TROPOMI data set provides a better support for the monitoring and controlling the VOCs related pollution.

310 *Data availability.* The data used in this paper is available on request from the corresponding author (chliu81@ustc.edu.cn).  
*Author contributions.* The first two authors contributed equally. CL and KLC designed and supervised the study and CL also help retrieve the satellite data. WS prepared and analyzed satellite data and wrote the manuscript. QH provided useful comments regarding the results. HL and XJ contributed to providing MAX-DOAS HCHO measurements. YZ and TL help ran the WRF-Chem model. CZ and YC help prepare satellite retrieval. JL supported the project.

315 *Acknowledgments.* This work was supported by grants from the National Key Research and Development Program of China (No. 2018YFC0213104, 2017YFC0210002 and 2016YFC0203302), Anhui Science and Technology Major Project (18030801111), National Natural Science Foundation of China (No. 41722501, 51778596 and 41977184), the Strategic Priority Research Program of the Chinese Academy of Sciences (No. XDA23020301), the National Key Project for Causes and Control of Heavy Air Pollution (No. DQGG0102 and DQGG0205), the National High-Resolution Earth Observation Project of China  
320 (No. 05-Y20A16-9001-15/17-2), Natural Science Foundation of Guangzhou Province (No.2016A030310115) and Civil Aerospace Technology Advance Research Project (No.Y7K00100KJ). The TROPOMI operational Level-1b product and HCHO product are available at <https://scihub.copernicus.eu/>.

*Competing interests.* The authors declare that they have no conflict of interest.



## References

- 325 Barkley, M. P., De Smedt, I., Van Roozendaal, M., Kurosu, T. P., Chance, K., Armeth, A., Hagberg, D., Guenther, A., Paulot, F., and Marais, E.: Top-down isoprene emissions over tropical South America inferred from SCIAMACHY and OMI formaldehyde columns, *Journal of Geophysical Research: Atmospheres*, 118, 6849-6868, <https://doi.org/10.1002/jgrd.50552>, 2013.
- Bauwens, M., Stavrakou, T., Müller, J.-F., Smedt, I. D., Roozendaal, M. V., Werf, G. R., Wiedinmyer, C., Kaiser, J. W.,
- 330 Sindelarova, K., and Guenther, A.: Nine years of global hydrocarbon emissions based on source inversion of OMI formaldehyde observations, *Atmos. Chem. Phys.*, 16, 10133-10158, <https://doi.org/10.5194/acp-16-10133-2016>, 2016.
- Bovensmann, H., Burrows, J., Buchwitz, M., Frerick, J., Noël, S., Rozanov, V., Chance, K., and Goede, A.: SCIAMACHY: Mission objectives and measurement modes, *J Atmos Sci*, 56, 127-150, [https://doi.org/10.1175/1520-0469\(1999\)056<0127:SMOAMM>2.0.CO;2](https://doi.org/10.1175/1520-0469(1999)056<0127:SMOAMM>2.0.CO;2), 1999.
- 335 Chan, K., Wiegner, M., Wenig, M., and Pöhler, D.: Observations of tropospheric aerosols and NO<sub>2</sub> in Hong Kong over 5 years using ground based MAX-DOAS, *Science of the total environment*, 619, 1545-1556, <https://doi.org/10.1016/j.scitotenv.2017.10.153>, 2018.
- Chan, K. L., Wang, Z., Ding, A., Heue, K.-P., Shen, Y., Wang, J., Zhang, F., Shi, Y., Hao, N., and Wenig, M.: MAX-DOAS measurements of tropospheric NO<sub>2</sub> and HCHO in Nanjing and a comparison to ozone monitoring instrument observations,
- 340 *Atmos. Chem. Phys.*, 19, 10051-10071, <https://doi.org/10.5194/acp-19-10051-2019>, 2019.
- Chance, K.: Analysis of BrO measurements from the global ozone monitoring experiment, *Geophys Res Lett*, 25, 3335-3338, <https://doi.org/10.1029/98GL52359>, 1998.
- Chance, K., and Kurosu, T.: SAO Participation in the GOME and SCIAMACHY Satellite Instrument Programs, 2003.
- Chance, K., and Kurucz, R. L.: An improved high-resolution solar reference spectrum for earth's atmosphere measurements in
- 345 the ultraviolet, visible, and near infrared, *Journal of quantitative spectroscopy and radiative transfer*, 111, 1289-1295, <https://doi.org/10.1016/j.jqsrt.2010.01.036>, 2010.
- Chance, K., and Orphal, J.: Revised ultraviolet absorption cross sections of H<sub>2</sub>CO for the HITRAN database, *Journal of Quantitative Spectroscopy and Radiative Transfer*, 112, 1509-1510, <https://doi.org/10.1016/j.jqsrt.2011.02.002>, 2011.



- Choi, Y., Kim, H., Tong, D., and Lee, P.: Summertime weekly cycles of observed and modeled NO<sub>x</sub> and O<sub>3</sub> concentrations as  
350 a function of satellite-derived ozone production sensitivity and land use types over the Continental United States, *Atmos.  
Chem. Phys.*, 12, 6291-6307, <https://doi.org/10.5194/acp-12-6291-2012>, 2012.
- De Smedt, I., Van Roozendael, M., Stavrou, T., Müller, J., Lerot, C., Theys, N., Valks, P., Hao, N., and Van Der A, R.:  
Improved retrieval of global tropospheric formaldehyde columns from GOME-2/MetOp-A addressing noise reduction and  
instrumental degradation issues, *Atmos. Meas. Tech.*, 5, 2933-2949, <https://doi.org/10.5194/amt-5-2933-2012>, 2012.
- 355 De Smedt, I., Theys, N., Yu, H., Danckaert, T., Lerot, C., Compennolle, S., Van Roozendael, M., Richter, A., Hilboll, A., and  
Peters, E.: Algorithm theoretical baseline for formaldehyde retrievals from S5P TROPOMI and from the QA4ECV project,  
<https://doi.org/10.5194/amt-11-2395-2018>, 2018.
- González Abad, G., Liu, X., Chance, K., Wang, H., Kurosu, T. P., and Suleiman, R.: Updated Smithsonian Astrophysical  
Observatory Ozone Monitoring Instrument (SAO OMI) formaldehyde retrieval, *Atmos. Meas. Tech.*, 8, 19-32,  
360 <https://doi.org/10.5194/amt-8-19-2015>, 2015.
- González Abad, G., Vasilkov, A., Seftor, C., Liu, X., and Chance, K.: Smithsonian Astrophysical Observatory Ozone Mapping  
and Profiler Suite (SAO OMPS) formaldehyde retrieval, *Atmos. Meas. Tech.*, 9, 2797-2812, <https://doi.org/10.5194/amt-9-2797-2016>, 2016.
- Guenther, A., Hewitt, C. N., Erickson, D., Fall, R., Geron, C., Graedel, T., Harley, P., Klinger, L., Lerdau, M., and McKay,  
365 W.: A global model of natural volatile organic compound emissions, *Journal of Geophysical Research: Atmospheres*, 100,  
8873-8892, <https://doi.org/10.1029/94JD02950>, 1995.
- Guenther, A., Karl, T., Harley, P., Wiedinmyer, C., Palmer, P., and Geron, C.: Estimates of global terrestrial isoprene emissions  
using MEGAN (Model of Emissions of Gases and Aerosols from Nature), *Atmos. Chem. Phys.*, 6, 3181-3210,  
<https://doi.org/10.5194/acp-6-3181-2006>, 2006.
- 370 Jang, M., and Kamens, R. M.: Characterization of secondary aerosol from the photooxidation of toluene in the presence of NO  
x and 1-propene, *Environ Sci Technol*, 35, 3626-3639, <https://doi.org/10.1021/es010676+>, 2001.



- Jiang, Z., Zheng, X., Zhai, H., Wang, Y., Wang, Q., and Yang, Z.: Seasonal and diurnal characteristics of carbonyls in the urban atmosphere of Changsha, a mountainous city in south-central China, *Environ Pollut*, 253, 259-267, <https://doi.org/10.1016/j.envpol.2019.06.127>, 2019.
- 375 Jin, X., and Holloway, T.: Spatial and temporal variability of ozone sensitivity over China observed from the Ozone Monitoring Instrument, *Journal of Geophysical Research: Atmospheres*, 120, 7229-7246, <https://doi.org/10.1002/2015JD023250>, 2015.
- Kraus, S.: DOASIS: A framework design for DOAS, Ph.D. thesis, University of Heidelberg, Germany, 2006.
- Li, H., He, Q., Song, Q., Chen, L., Song, Y., Wang, Y., Lin, K., Xu, Z., and Shao, M.: Diagnosing Tibetan pollutant sources via volatile organic compound observations, *Atmos Environ*, 166, 244-254, <https://doi.org/10.1016/j.atmosenv.2017.07.031>,  
380 2017.
- Li, J., Wang, Z., and Xiang, W.: Daytime atmospheric oxidation capacity of urban Beijing under polluted conditions during the 2008 Beijing Olympic Games and the impact of aerosols, *Sola*, 7, 73-76, <https://doi.org/10.2151/sola.2011-019>, 2011.
- Li, L., and Xie, S.: Historical variations of biogenic volatile organic compound emission inventories in China, 1981–2003, *Atmos Environ*, 95, 185-196, <https://doi.org/10.1016/j.atmosenv.2014.06.033>, 2014.
- 385 Liu, H., Liu, C., Xie, Z., Li, Y., Huang, X., Wang, S., Xu, J., and Xie, P.: A paradox for air pollution controlling in China revealed by "APEC Blue" and "Parade Blue", *Sci Rep*, 6, 34408, <https://doi.org/10.1038/srep34408>, 2016.
- Loyola, D. G., Gimeno García, S., Lutz, R., Argyrouli, A., Romahn, F., Spurr, R. J., Pedernana, M., Doicu, A., Molina García, V., and Schüssler, O.: The operational cloud retrieval algorithms from TROPOMI on board Sentinel-5 Precursor, *Atmos. Meas. Tech.*, 11, <https://doi.org/10.5194/amt-11-409-2018>, 2018.
- 390 Malicet, J., Daumont, D., Charbonnier, J., Parrisé, C., Chakir, A., and Brion, J.: Ozone UV spectroscopy. II. Absorption cross-sections and temperature dependence, *J Atmos Chem*, 21, 263-273, <https://doi.org/10.1007/BF00696758>, 1995.
- Martin, R., Parrish, D., Ryerson, T., Nicks, D., Chance, K., Kurosu, T., Jacob, D. J., Sturges, E., Fried, A., and Wert, B.: Evaluation of GOME satellite measurements of tropospheric NO<sub>2</sub> and HCHO using regional data from aircraft campaigns in the southeastern United States, *Journal of Geophysical Research: Atmospheres*, 109, <https://doi.org/10.1029/2004jd004869>,  
395 2004.



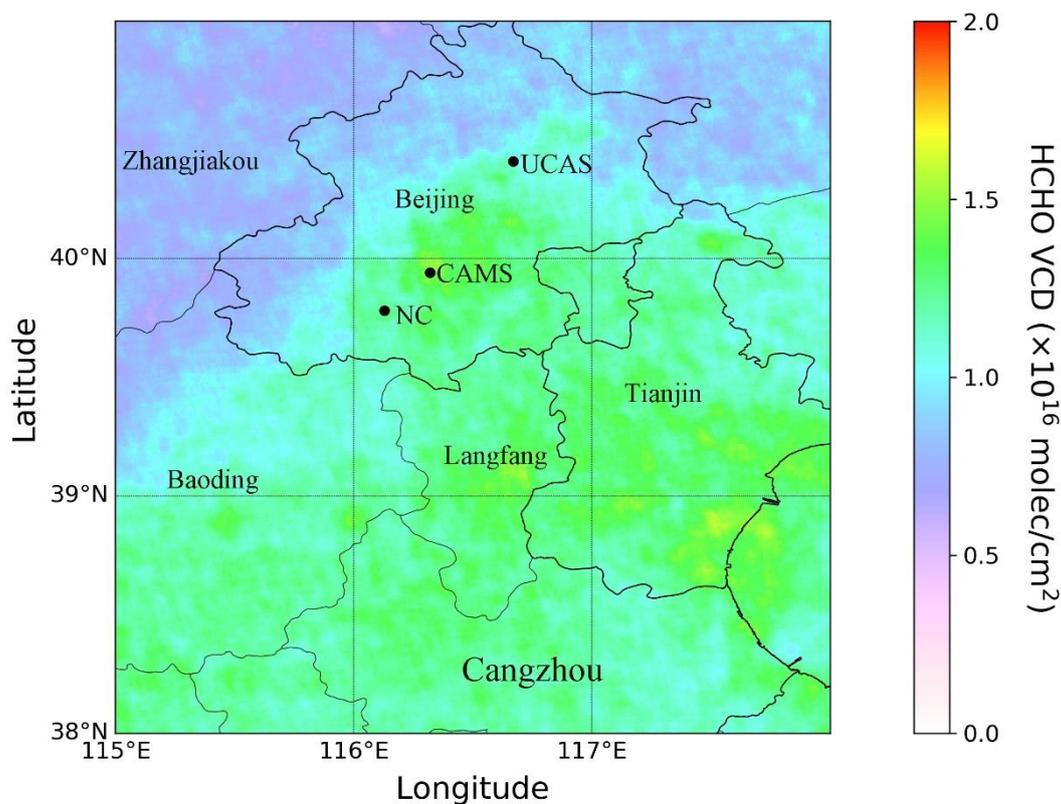
- Martin, R. V., Chance, K., Jacob, D. J., Kurosu, T. P., Spurr, R. J., Bucsela, E., Gleason, J. F., Palmer, P. I., Bey, I., and Fiore, A. M.: An improved retrieval of tropospheric nitrogen dioxide from GOME, *Journal of Geophysical Research: Atmospheres*, 107, ACH 9-1-ACH 9-21, <https://doi.org/10.1029/2001JD001027>, 2002.
- Palmer, P. I., Jacob, D. J., Chance, K., Martin, R. V., Spurr, R. J., Kurosu, T. P., Bey, I., Yantosca, R., Fiore, A., and Li, Q.:  
400 Air mass factor formulation for spectroscopic measurements from satellites: Application to formaldehyde retrievals from the Global Ozone Monitoring Experiment, *Journal of Geophysical Research: Atmospheres*, 106, 14539-14550, <https://doi.org/10.1029/2000JD900772>, 2001.
- Rodgers, C. D.: *Inverse methods for atmospheric sounding: theory and practice*, World scientific, 2000.
- Schönhardt A, R. A., Wittrock F: Observations of iodine monoxide columns from satellite, *Atmos. Chem. Phys.*,  
405 <https://doi.org/10.5194/acp-8-637-2008>, 2008.
- Su, W., Liu, C., Hu, Q., Zhao, S., Sun, Y., Wang, W., Zhu, Y., Liu, J., and Kim, J.: Primary and secondary sources of ambient formaldehyde in the Yangtze River Delta based on Ozone Mapping and Profiler Suite (OMPS) observations, *Atmos. Chem. Phys.*, 19, 6717-6736, <https://doi.org/10.5194/acp-19-6717-2019>, 2019.
- Thalman, R., and Volkamer, R.: Temperature dependent absorption cross-sections of O<sub>2</sub>-O<sub>2</sub> collision pairs between 340 and  
410 630 nm and at atmospherically relevant pressure, *Phys Chem Chem Phys*, 15, 15371-15381, <https://doi.org/10.1039/c3cp23017a>, 2013.
- Vandaele, A. C., Hermans, C., Simon, P. C., Carleer, M., Colin, R., Fally, S., Merienne, M.-F., Jenouvrier, A., and Coquart, B.: Measurements of the NO<sub>2</sub> absorption cross-section from 42000 cm<sup>-1</sup> to 10000 cm<sup>-1</sup> (238–1000 nm) at 220 K and 294 K, *Journal of Quantitative Spectroscopy and Radiative Transfer*, 59, 171-184, [https://doi.org/10.1016/S0022-4073\(97\)00168-4](https://doi.org/10.1016/S0022-4073(97)00168-4), 1998.
- 415 Wang, D., Zhou, B., Fu, Q., Zhao, Q., Zhang, Q., Chen, J., Yang, X., Duan, Y., and Li, J.: Intense secondary aerosol formation due to strong atmospheric photochemical reactions in summer: observations at a rural site in eastern Yangtze River Delta of China, *Science of the Total Environment*, 571, 1454-1466, <https://doi.org/10.1016/j.scitotenv.2016.06.212>, 2016.
- Wei, W., Wang, S., Chatani, S., Klimont, Z., Cofala, J., and Hao, J.: Emission and speciation of non-methane volatile organic compounds from anthropogenic sources in China, *Atmos Environ*, 42, 4976-4988,  
420 <https://doi.org/10.1016/j.atmosenv.2008.02.044>, 2008.



Wilmouth, D. M., Hanisco, T. F., Donahue, N. M., and Anderson, J. G.: Fourier transform ultraviolet spectroscopy of the  $A\ 2\Pi_{3/2} \leftarrow X\ 2\Pi_{3/2}$  transition of BrO, *The journal of physical chemistry a*, 103, 8935-8945, <https://doi.org/10.1021/jp991651o>, 1999.

425 Wu, R., Bo, Y., Li, J., Li, L., Li, Y., and Xie, S.: Method to establish the emission inventory of anthropogenic volatile organic compounds in China and its application in the period 2008–2012, *Atmos Environ*, 127, 244-254, <https://doi.org/10.1016/j.atmosenv.2015.12.015>, 2016.

Xue, L., Gu, R., Wang, T., Wang, X., Saunders, S., Blake, D., Louie, P. K., Luk, C. W., Simpson, I., and Xu, Z.: Oxidative capacity and radical chemistry in the polluted atmosphere of Hong Kong and Pearl River Delta region: analysis of a severe photochemical smog episode, *Atmos. Chem. Phys.*, <https://doi.org/10.5194/acp-16-9891-2016>, 2016.



430

**Figure1:** Annual average of the improved TROPOMI tropospheric HCHO VCD from August 2018 to July 2019 in Beijing and its surrounding region. The black dots indicate the locations of MAX-DOAS sites.

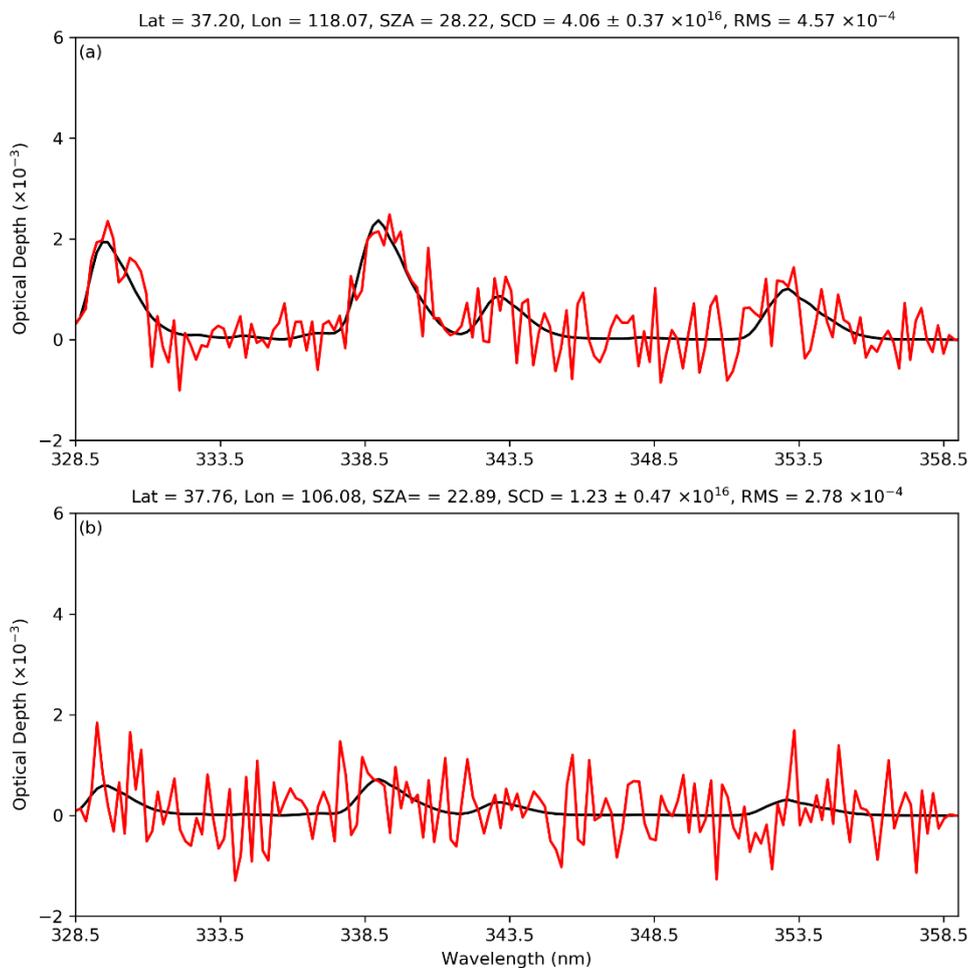
**Table 1.** Information of the MAX-DOAS measurements.



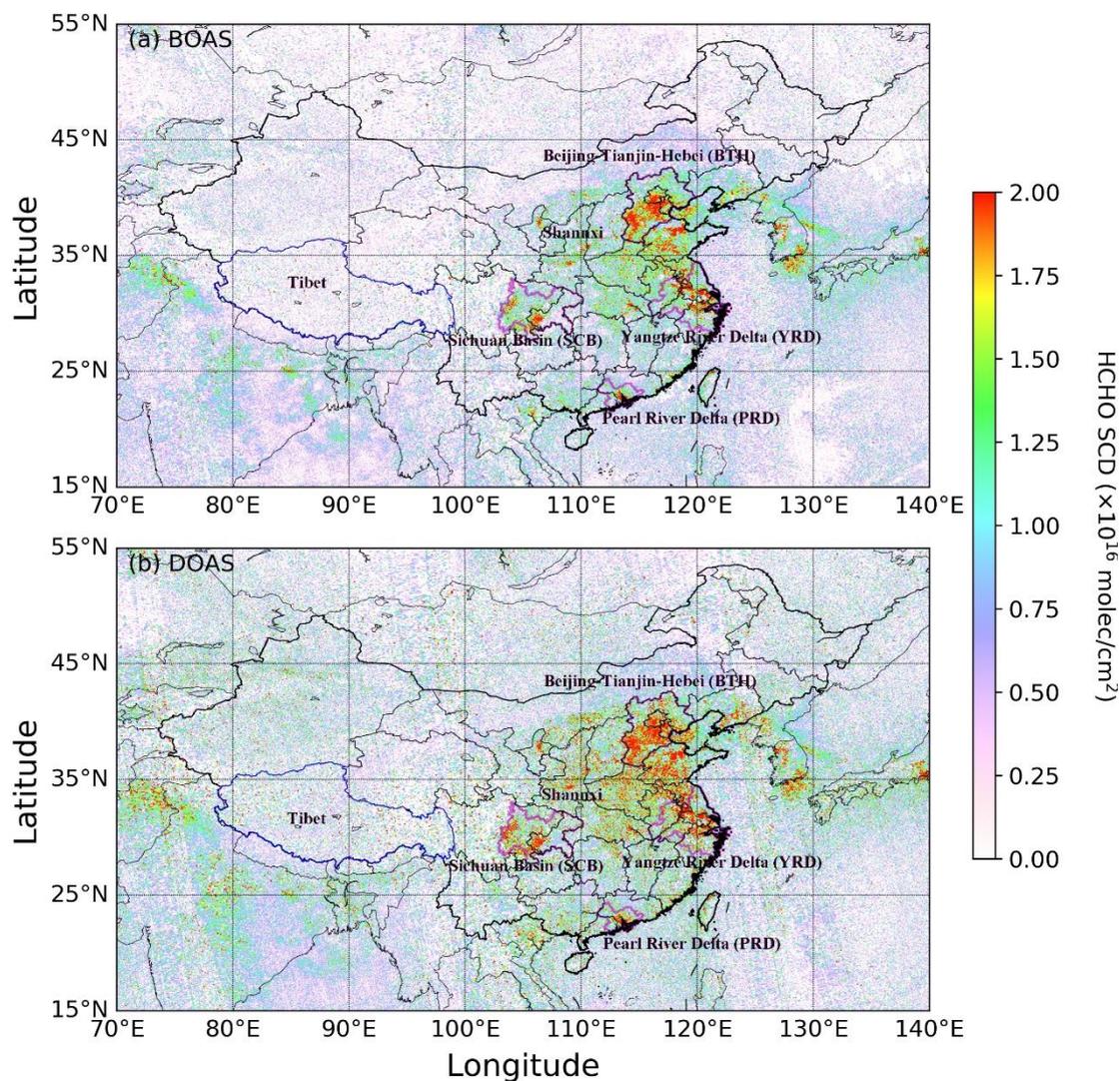
| Site name | Location               | Region     | Height<br>above sea<br>level (m) | viewing azimuth<br>angle (°) | elevation<br>angles (°)                |
|-----------|------------------------|------------|----------------------------------|------------------------------|--|
| CAMS      | 39.94° N,<br>116.32° E | urban area | 100                              | 130 (southeast)              | 1, 2, 3, 4, 5, 6,<br>8, 10, 15, 30, 90 |
| UCAS      | 40.41° N,<br>116.67° E | suburb     | 120                              | 67 (northeast)               | 1, 2, 3, 4, 5, 6,<br>8, 10, 15, 30, 90 |
| NC        | 39.78° N,<br>116.13° E | suburb     | 60                               | 48 (northeast)               | 1, 2, 3, 4, 5, 6,<br>8, 10, 15, 30, 90 |

435 **Table 2: The retrieval settings for TROPOMI HCHO DSCD in our retrieval.**

|                                |   |
|--------------------------------|---|
| Fitting window                 | 328.5-359 nm  |
| Radiance reference spectrum    | Computed online from the Pacific orbit between<br>30°S and 30°N |
| Scaling polynomial             | 3rd order   |
| Baseline polynomial            | 3rd order   |
| Instrument slit function       | Pre-flight measurements   |
| Solar reference spectrum       | (Chance and Kurucz, 2010)                                       |
| HCHO cross sections            | (Chance and Orphal, 2011) , 300K                                |
| O <sub>3</sub> cross sections  | (Malicet et al., 1995) , 228 and 295K                           |
| NO <sub>2</sub> cross sections | (Vandaele et al., 1998) , 220K                                  |
| BrO cross sections             | (Wilmouth et al., 1999) , 228K                                  |
| O <sub>4</sub> cross sections  | (Thalman and Volkamer, 2013) , 293K                             |

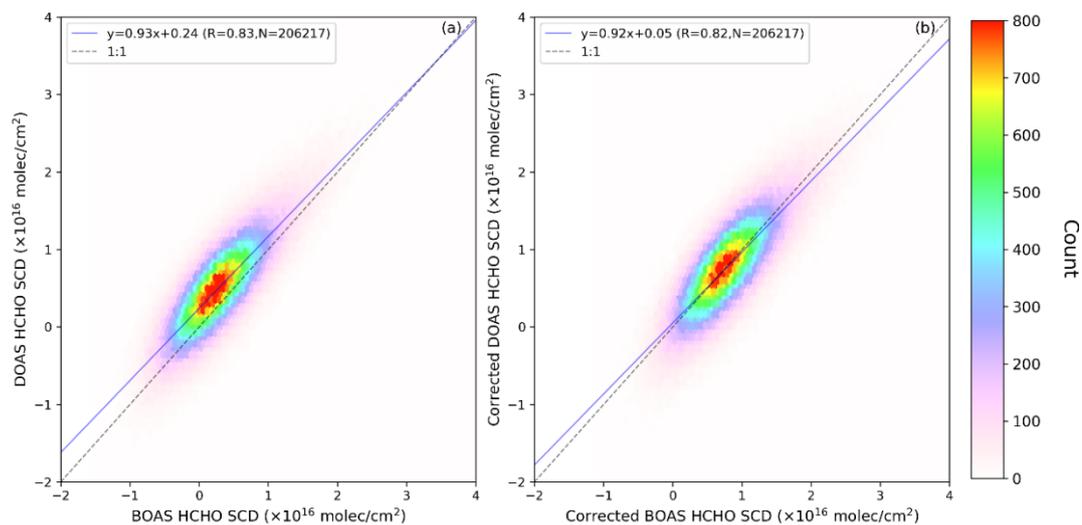


**Figure 2:** An example of the spectral retrieval of HCHO SCD from two TROPOMI spectra measured on 6 August 2019 over China (orbit 4211), for (a) polluted case and (b) clean case. The black lines represent the simulated HCHO optical depth, and the red lines represent the fitted HCHO optical depth plus fitting residuals.



440

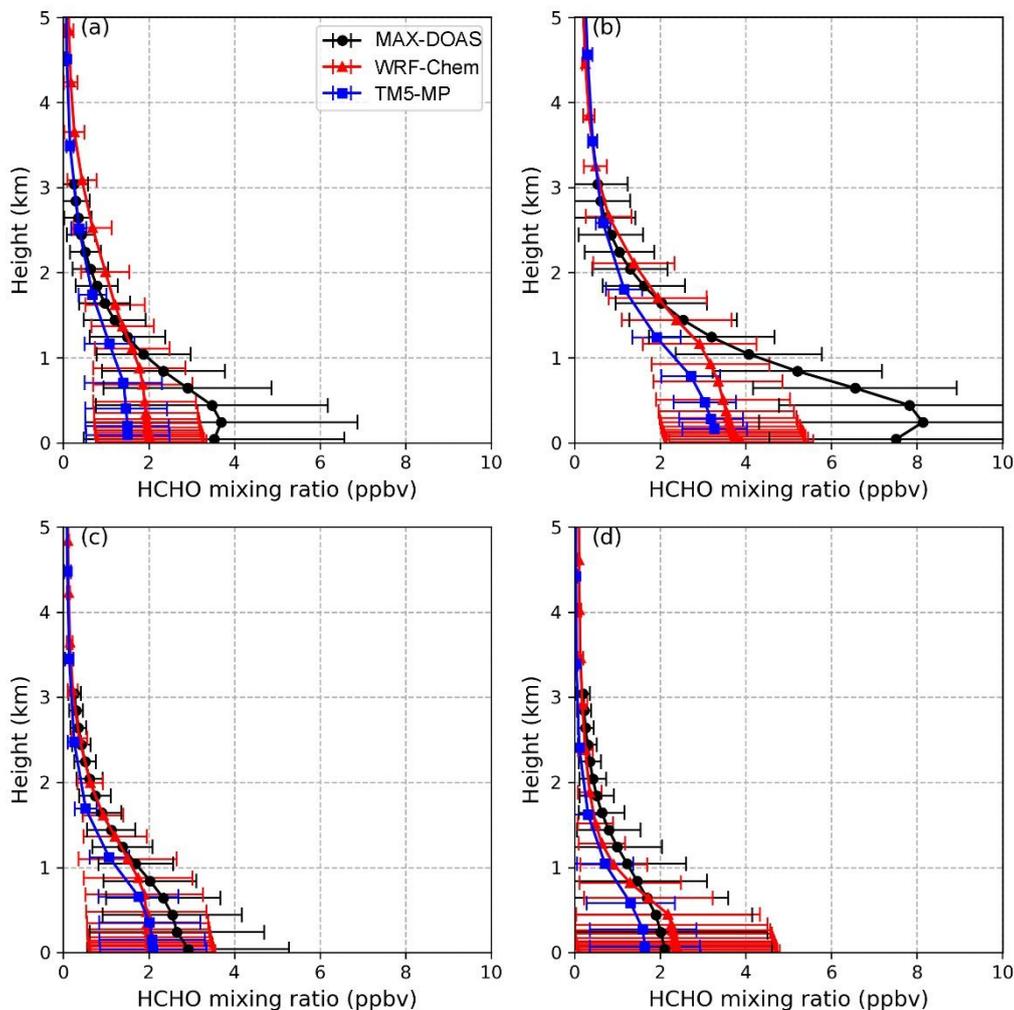
**Figure 3: Spatial distributions of SCDs retrieved by BOAS (a) and DOAS (b) methods on 06 August 2018 over China. The regional boundaries of BTH, YRD, PRD and SCB are delineated by magenta lines.**



445 **Figure 4:** (a) Pixel to pixel comparisons of DOAS HCHO SCDs and BOAS HCHO SCDs and (b) corrected DOAS HCHO SCDs and corrected BOAS HCHO SCDs on 06 August 2018 in the region between 73° E and 130° E, and 18° N and 54° N.

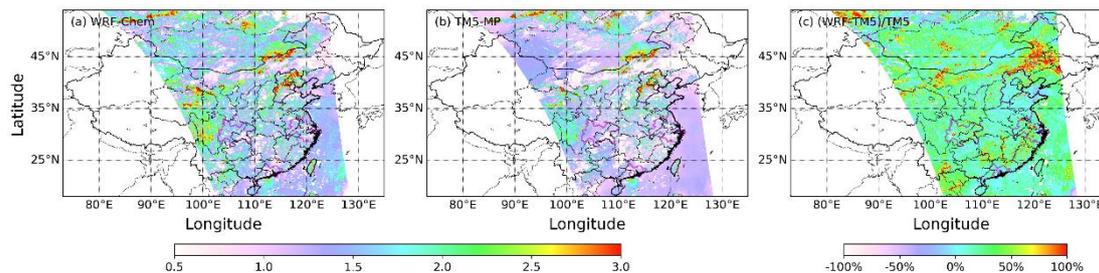
**Table 3:** NMBs between satellite and MAX-DOAS observations. TROPOMI HCHO VCDs with different retrieval settings are shown. HCHO VCDs are calculated with 3 different settings, (1) by replacing SCD retrieved with BOAS in the operational product, (2) changing the a priori profiles from TM5 to regional WRF-Chem simulations and (3) both (1) and (2) changes.

| Settings (S1 vs S2)                | CAMS                 |                     | UCAS                 |                     | NC                   |                     |
|------------------------------------|----------------------|---------------------|----------------------|---------------------|----------------------|---------------------|
|                                    | NMB <sub>S1,S2</sub> | NMB <sub>S1,M</sub> | NMB <sub>S1,S2</sub> | NMB <sub>S1,M</sub> | NMB <sub>S1,S2</sub> | NMB <sub>S1,M</sub> |
| Operational product                | 0 %                  | 7.32 %              | 0 %                  | 23.86 %             | 0 %                  | 10.84 %             |
| BOAS vs DOAS                       | 0.20 %               | 7.47 %              | -0.13 %              | 23.69 %             | -0.15 %              | 10.68 %             |
| WRF-Chem vs TM5-MP                 | -9.83 %              | -3.24 %             | -20.33 %             | -1.32 %             | -16.32 %             | -7.24 %             |
| BOAS vs DOAS<br>WRF-Chem vs TM5-MP | -8.81 %              | -2.14 %             | -20.81 %             | -1.91 %             | -16.06 %             | -6.95 %             |



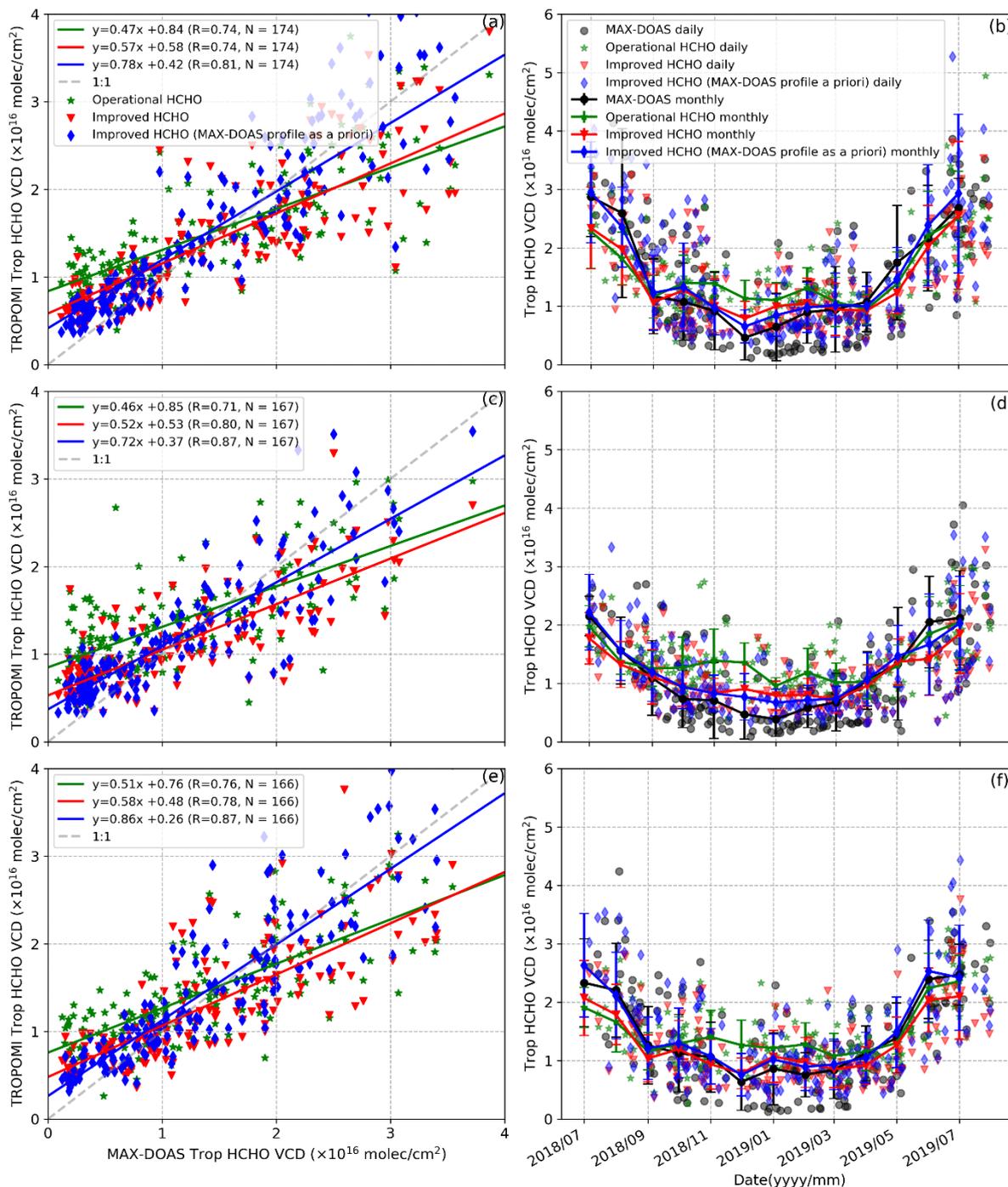
450

**Figure 5: Vertical HCHO profiles obtained from MAX-DOAS, WRF-Chem and TM5-MP model in spring (a), summer (b), autumn (c) and winter (d) at CAMS site. The error bars represent 1 $\sigma$  standard deviation of variation.**

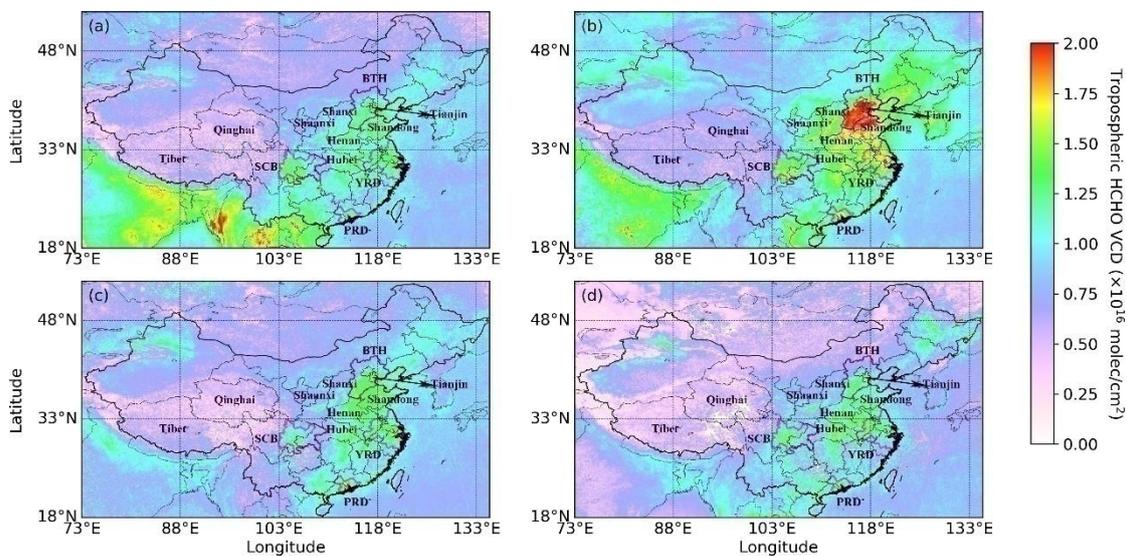


455

**Figure 6: Spatial distributions of AMFs calculated using WRF-Chem (a) and TM5-MP (b) simulations as a priori HCHO profiles of orbit 04211 on 06 August 2018. (c) Spatial distribution of difference of the AMFs.**



**Figure 7: Comparison of tropospheric HCHO VCDs between TROPOMI and MAX-DOAS measurements from July 2018 to July 2019 at CAMS (a), UCAS (c) and NC (e) sites. Time series of TROPOMI and MAX-DOAS tropospheric HCHO VCDs from July 2018 to July 2019 at CAMS (b), UCAS (d) and NC (f) sites.**



460

**Figure 8:** The spatial distribution of average tropospheric HCHO VCDs over China in spring (a), summer (b), winter (c) and winter (d).